Сильно динамически устойчивые решения в динамических кооперативных играх

Л. А. Петросян

СПбГУ

1-3 апреля 2015
Екатеринбург

Постановка задачи

Рассматривается кооперативная дифференциальная игра n лиц $\Gamma(x_0, T - t_0)$ с предписанной продолжительностью $T - t_0$, на отрезке времени $t \in [t_0, T]$ из начального состояния $x_0 \in R^n$, уравнениями движения

$$\dot{x} = f(x, u_1, \ldots, u_n), \qquad x(t_0) = x_0, \tag{1}$$

$x \in R^n$, $u_i \in U \subset compR^k$, и функциями выигрыша

$$K_i(x_0, T - t_0; u_1, \ldots, u_n) = \int_{t_0}^{T} h_i(x(t))dt,$$

где $x(t)$ — решение системы (1) при управлениях $(u_1, \ldots, u_n)$, и $i \in N$, где N — множество игроков и $|N| = n$.

Пусть $S \subset N$ — коалиция в игре $\Gamma(x_0, T - t_0)$. Определим для $S \subset N$ характеристическую функцию $v(x_0, T - t_0; S)$ как нижнее значение антагонистической игры между коалицией S, действующей как игрок I (максимизирующий) и коалицией $N \backslash S$, действующей как игрок II (минимизирующий), где под выигрышем игрока S понимается сумма выигрышей игроков, входящих в S, а под стратегией игрока S — элемент декартова произведения множеств стратегий игроков, входящих в S. В данной формализации, мы из соображений простоты под стратегией игрока будем понимать функцию $u_i(x, t)$ со значением в множестве мгновенных допустимых управлений U_i. Нижнее значение игры (supinf) всегда существует и является супераддитивной функцией от коалиции $S \subset N$.

Кооперативная траектория

Определение 1.

Траекторию $\bar{x}(t)$, $t \in [t_0, T]$ назовем кооперативной траекторией, если имеет место

$$\max_{u_1,\ldots,u_n} \sum_{i=1}^{n} K_i(x_0, T - t_0; u_1, \ldots, u_n) =$$
$$= \max_{x(t)} \sum_{i=1}^{n} \int_{t_0}^{T} h_i(x(t))dt = \sum_{i=1}^{n} \int_{t_0}^{T} h_i(\bar{x}(t)).$$

Мы предполагаем, что такая траектория $\bar{x}(t)$, $t \in [t_0, T]$ существует. В противном случае, дальнейшие выкладки и определения необходимо незначительно изменить.

Рассмотрим подыгры игры $\Gamma(x_0, T - t_0)$, $\Gamma(\bar{x}(t), T - t)$ с начальным условием на кооперативной траектории. В каждой подыгре $\Gamma(\bar{x}(t), T - t)$ можно таким же образом, как и в основной игре $\Gamma(x_0, T - t_0)$ определить характеристическую функцию (Воробьев, 1981) $v(\bar{x}(t), T - t; S)$, где $S \subset N$, которая также будет супераддитивной. Определим множество дележей $M(\bar{x}(t), T - t)$ в игре $\Gamma(\bar{x}(t), T - t)$ как

$$M(\bar{x}(t), T - t) = \{\alpha = (\alpha_1, \dots, \alpha_n) : \sum_{i=1}^{n} \alpha_i = v(\bar{x}(t), T - t; N),$$
$$\alpha_i \geqslant v(\bar{x}(t), T - t; \{i\}), i \in N\}.$$

Из супераддитивности характеристической функции следует, что множество $M(\bar{x}(t), T - t) \neq \emptyset$, $t \in [t_0, T]$. Определим также ядро $C(\bar{x}(t), T - t) \subset M(\bar{x}(t), T - t)$ в игре $\Gamma(\bar{x}(t), T - t)$ и предположим, что для всех $t \in [t_0, T]$, $C(\bar{x}(t), T - t) \neq \emptyset$.

Процедура распределения дележа

Напомним, что ядро в игре $\Gamma(\bar{x}(t), T-t)$ — это множество дележей $\alpha^t = (\alpha_1^t, \ldots, \alpha_n^t)$, удовлетворяющих неравенствам

$$\sum_{i \in S} \alpha_i^t \geqslant v(\bar{x}(t), T-t; S)$$

для всех $S \subset N$.

Определение 2

(Петросян и Данилов, 1979) Функция $\beta_i(\tau)$, $\tau \in [t_0, T]$, $i \in N$ называется процедурой распределения дележа $\alpha \in M(x_0, T-t_0)$, если

$$\alpha_i = \int_{t_0}^{T} \beta_i(\tau) d\tau, \qquad i \in N.$$

Динамически устойчивое ядро

Определение 3

(Петросян и Данилов, 1979) Ядро $C(x_0, T - t_0)$ в игре $\Gamma(x_0, T - t_0)$ называется динамически устойчивым, если для каждого дележа $\alpha \in C(x_0, T - t_0)$ найдется процедура распределения дележа (ПРД) такая, что

$$\int_t^T \beta_i(\tau) d\tau \in C(\bar{x}(t), T - t), \qquad t \in [t_0, T], i \in N.$$

Нами было ранее показано (Петросян и Данилов, 1979; Yeung&Petrosjan, 2006), что если $C(\bar{x}(t), T-t) \neq \emptyset$ при $t \in [t_0, T]$, и существует дифференцируемый селектор $\alpha(t) \in C(\bar{x}(t), T-t)$ ($\alpha(t_0) = \alpha$), то ядро является динамически устойчивым и ПРД $\beta_i(T)$ определяется по формуле

$$\begin{aligned} \beta_i(\tau) &= -\frac{d}{d\tau}\alpha(\tau), \\ \alpha(t_0) &= \alpha. \end{aligned}$$

Приведем определение сильного динамически устойчивого ядра $C(x_0, T-t_0)$ в игре $\Gamma(x_0, T-t_0)$.

Сильно динамически устойчивое ядро

Определение 4

Ядро $C(x_0, T - t_0)$ сильно динамически устойчиво в игре $\Gamma(x_0, T - t_0)$, если

1. $C(\bar{x}(t), T - t) \neq \emptyset$, $t \in [t_0, T]$.
2. Существует такой дележ $\alpha \in C(x_0, T - t_0)$ и такая ПРД $\beta(\tau) = (\beta_1(\tau), \dots, \beta_n(\tau))$, $\tau \in [t_0, T]$, что $\alpha = \displaystyle\int_{t_0}^{T} \beta_i(\tau) d\tau$ и

$$C(x_0, T - t_0) \supset \int_{t_0}^{t} \beta(\tau) d\tau \oplus C(\bar{x}(t), T - t), \ \ t \in [t_0, T].$$

Здесь символ $\oplus$ определяется следующим образом. Пусть $a \in R^n$, $B \subset R^n$, тогда $a \oplus B = \{a + b : b \in B\}$.
Заметим, что Определение 4 незначительно отличается от определения сильной динамической устойчивости в работах Петросян, 1995a; Петросян, 1995b.

Как мы уже отметили и как это следует непосредственно из определения множества дележей и ядра для игры двух лиц, эти множества совпадают, т.е. $M(x_0, T - t_0) = C(x_0, T - t_0)$, и также множества $M(\bar{x}(t), T - t) = C(\bar{x}(t), T - t)$, $t \in [t_0, T]$.
Для игры $\Gamma(x_0, T - t_0)$ с двумя участниками ядро имеет вид

$$C(x_0, T - t_0) = \{\alpha = (\alpha_1, \alpha_2) : \ \alpha_1 + \alpha_2 = v(x_0, T - t_0; N),$$

$$\alpha_1 \geq v(x_0, T - t_0; \{1\}), \alpha_2 \geq v(x_0, T - t_0; \{2\})\} = M(x_0, T - t_0).$$

Аналогичным образом для подыгр $\Gamma(x(t), T - t)$ мы имеем

$$C(\bar{x}(t), T - t) = \{\alpha^t = (\alpha_1^t, \alpha_2^t) : \ \alpha_1^t + \alpha_2^t = v(\bar{x}(t), T - t; N),$$

$$\alpha_1^t \geq v(\bar{x}(t), T - t; \{1\}), \alpha_2^t \geq v(\bar{x}(t), T - t; \{2\})\} = M(\bar{x}(t), T - t).$$

Введем величины $A_1 \geq 0$ и $A_2 \geq 0$ следующим образом

$$A_1 + A_2 = v(x_0, T - t_0; N) - \sum_{i=1}^{2} v(x_0, T - t_0; \{i\}) \geq 0 \qquad (2)$$

Представим A_i в виде $A_i = \int\limits_{t_0}^{T} \gamma_i(\tau) d\tau$, $i = 1, 2$.

Введем функции (в предположении дифференцируемости функций $v(\bar{x}(t), T - t; \{i\})$)

$$\bar{\beta}_i(\tau) = \gamma_i(\tau) - \frac{d}{d\tau} v(\bar{x}(\tau), T - \tau; \{i\}))$$

и потребуем, чтобы при всех $\tau \in [t_0, T]$ имело место

$$\bar{\beta}_1(\tau) + \bar{\beta}_2(\tau) = -\frac{d}{d\tau} v(\bar{x}(\tau), T - \tau; N)). \qquad (3)$$

Тогда получаем

$$\sum_{i=1}^{2} \gamma_i(\tau) = -\frac{d}{d\tau} v(\bar{x}(\tau), T-\tau; N) + \sum_{i=1}^{2} \frac{d}{d\tau} v(\bar{x}(\tau), T-\tau; \{i\}) \qquad (4)$$

и, интегрируя, получаем (2)

$$A_1 + A_2 = -\int_{t_0}^{T} \frac{d}{d\tau} v(\bar{x}(\tau), T-\tau; N) d\tau +$$

$$+ \sum_{i=1}^{2} \int_{t_0}^{T} \frac{d}{d\tau} v(\bar{x}(\tau), T-\tau; \{i\}) d\tau.$$

$$A_1 + A_2 = v(\bar{x}(t_0), T-t_0; N) - \sum_{i=1}^{2} v(\bar{x}(t_0), T-t_0; \{i\}) \geqslant 0.$$

То есть, если $\gamma_i(\tau)$, $i = 1, 2$ удовлетворяет (4), то условие (2) всегда выполнено.

Рассмотрим теперь вектор $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2)$, где

$$\bar{\alpha}_i = \int_{t_0}^{T} \bar{\beta}_i(\tau) d\tau \tag{5}$$

Легко видеть, что $\bar{\alpha} \in C(x_0, T - t_0)$. Действительно,

$$\bar{\alpha}_i = \int_{t_0}^{T} \bar{\beta}_i(\tau) d\tau = \int_{t_0}^{T} \gamma_i(\tau) d\tau + \int_{t_0}^{T} (-\frac{d}{d\tau} v(\bar{x}(\tau), T - \tau; \{i\})) d\tau =$$

$$= A_i + v(x_0, T - t_0; \{i\}),$$

и согласно (2)

$$\bar{\alpha}_1 + \bar{\alpha}_2 = \sum_{i=1}^{2} A_i + v(x_0, T - t_0; \{i\}) = v(x_0, T - t_0; N).$$

А поскольку $A_i > 0$, имеем

$$\bar{\alpha}_i = A_i + v(x_0, T - t_0; \{i\}) \geq v(x_0, T - t_0; \{i\}),$$

т.е. $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2) \in C(x_0, T - t_0)$.

Условие 2.1

Существуют такие A_1 и A_2, удовлетворяющие (2) и такие $\gamma_1(\tau)$, $\gamma_2(\tau)$, $\tau \in [t_0, T]$, удовлетворяющие (3), что $\int\limits_{t_0}^{t} \gamma_i(\tau) d\tau \geqslant 0$ при $t \in [t_0, T]$.

Теорема 1

При выполнении условия 2.1 ядро игры $\Gamma(x_0, T - t_0)$ сильно динамически устойчиво.

Заметим, что в каждый момент времени $t \in [t_0, T]$ имеет место

$$\sum_{i=1}^{2} \beta_i(t) = -\frac{d}{dt} v(\bar{x}(t), T - t; N) = \sum_{i=1}^{2} h_i(\bar{x}(t)),$$

т.е. используя ПРД $\bar{\beta} = (\bar{\beta}_1, \bar{\beta}_2)$ игроки фактически в каждый момент времени перераспределяют мгновенный суммарный доход, т.е. обеспечивается мгновенная трансферабельность выигрышей, что вполне соответствует идеологии теории кооперативных игр с трансферабельными выигрышами.

Случай трех игроков

В игре трех лиц возможны следующие коалиции: $\{1\}$, $\{2\}$, $\{3\}$, $\{1,2\}$, $\{2,3\}$, $\{1,3\}$, $\{1,2,3\} = N$.
Рассмотрим систему уравнений

$$\begin{aligned} \alpha_1 + \alpha_2 &= v(x_0, T - t_0; \{1,2\}) = a, \\ \alpha_2 + \alpha_3 &= v(x_0, T - t_0; \{2,3\}) = b, \\ \alpha_1 + \alpha_3 &= v(x_0, T - t_0; \{1,3\}) = c. \end{aligned} \tag{6}$$

Решение этой системы имеет вид

$$\alpha_1 = \frac{a-b+c}{2}, \quad \alpha_2 = \frac{b-c+a}{2}, \quad \alpha_3 = \frac{c+b-a}{2}.$$

Предположим, что имеет место

$$\alpha_1 + \alpha_2 + \alpha_3 \leq v(x_0, T - t_0; \{1,2,3\}) = v(x_0, T - t_0; N),$$

т.е.

$$\frac{a-b+c}{2} + \frac{b-c+a}{2} + \frac{c+b-a}{2} = \frac{a+b+c}{2} \leq v(x_0, T - t_0; N).$$

Рассмотрим теперь уравнение (6) на отрезке времени $[t_0, T]$, когда величины $v(\bar{x}(t), T-t; \{1,2\}) = a(t)$, $v(\bar{x}(t), T-t; \{2,3\}) = b(t)$, $v(\bar{x}(t), T-t; \{1,3\}) = c(t)$ вычисляются вдоль кооперативной траектории.

Тогда будем иметь

$$\begin{aligned} \alpha_1(t) + \alpha_2(t) &= v(\bar{x}(t), T-t; \{1,2\}) = a(t), \\ \alpha_2(t) + \alpha_3(t) &= v(\bar{x}(t), T-t; \{2,3\}) = b(t), \\ \alpha_1(t) + \alpha_3(t) &= v(\bar{x}(t), T-t; \{1,3\}) = c(t). \end{aligned} \tag{7}$$

И решение системы (7) будет иметь вид

$$\alpha_1(t) = \frac{a(t) - b(t) + c(t)}{2}, \alpha_2(t) = \frac{b(t) - c(t) + a(t)}{2},$$

$$\alpha_3(t) = \frac{c(t) + b(t) - a(t)}{2}.$$

Пусть

$$A(t) = v(\bar{x}(t), T - t; N) - \sum_{i=1}^{3} \alpha_i(t), \tag{8}$$

и $A_i(t)$ удовлетворяют условию

$$\sum_{i=1}^{3} A_i(t) = A(t),$$

Предполагая дифференцируемость функций $A_i(t)$, положим

$$\gamma_i(t) = -\frac{d}{dt} A_i(t).$$

Условие 3.1

Предположим, что можно выбрать $A_i(t)$ таким образом, чтобы $\gamma_i(t) \geq 0$, и $\alpha_i(t) \geqslant v(\bar{x}(t), T - t; \{i\})$, $t \in [t_0, T]$, $i = 1, 2, 3$.

Введем функции

$$\bar{\beta}_1(t) = \gamma_1(t) - \frac{d}{dt}\bar{\alpha}_1(t),$$

$$\bar{\beta}_2(t) = \gamma_2(t) - \frac{d}{dt}\bar{\alpha}_2(t),$$

$$\bar{\beta}_3(t) = \gamma_3(t) - \frac{d}{dt}\bar{\alpha}_3(t).$$

Положим

$$\int_{t_0}^{T} \bar{\beta}_i(\tau)d\tau = \bar{\alpha}_i \geq 0, \quad i \in N,$$

и будем иметь

$$\sum_{i=1}^{3} \bar{\alpha}_i = v(x_0, T - t_0; N), \quad \bar{\alpha}_i \geq v(x_0, T - t_0; \{i\}) \quad i \in N,$$

$$\bar{\alpha}_1 + \bar{\alpha}_2 \geq v(x_0, T - t_0; \{1, 2\}),$$

$$\bar{\alpha}_2 + \bar{\alpha}_3 \geq v(x_0, T - t_0; \{2, 3\}),$$

$$\bar{\alpha}_1 + \bar{\alpha}_3 \geq v(x_0, T - t_0; \{1, 3\}).$$

Таким образом, вектор $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3)$ принадлежит ядру $C(x_0, T - t_0)$ игры $\Gamma(x_0, T - t_0)$.

Теорема 2

При выполнении условия 3.1 имеет место сильная динамическая устойчивость ядра $C(x_0, T - t_0)$, при этом в качестве дележа $\alpha = (\alpha_1, \alpha_2, \alpha_3)$, фигурирующего в определении сильной динамической устойчивости, можно взять дележ $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3)$.

Заметим, что также как и в случае двух игроков в каждый момент времени $t \in [t_0, T]$ имеет место (см. (8))

$$v(\bar{x}(t), T - t; N) = \sum_{i=1}^{3} [A_i(t) + \alpha_i(t)]$$

или

$$-\frac{d}{dt} v(\bar{x}(t), T - t; N) = \sum_{i=1}^{3} \gamma_i(t) - \sum_{i=1}^{3} \frac{d}{dt} \alpha_i(t),$$

что дает нам

$$\sum_{i=1}^{3} h_i(\bar{x}(t)) = \sum_{i=1}^{3} [\gamma_i(t) - \frac{d}{dt} \alpha_i(t)] = \sum_{i=1}^{3} \bar{\beta}_i(t).$$

Таким образом, используя ПРД $\bar{\beta} = (\bar{\beta}_1, \bar{\beta}_2, \bar{\beta}_3)$ игроки в каждый момент времени перераспределяют мгновенный суммарный доход, т. е. обеспечивается мгновенная трансферабельность выигрышей, что вполне соответствует идеологии кооперативных игр с трансф. выигрышами.

Список литературы

Воробьев Н.Н. Теория игр для экономистов-кибернетиков. М.: Наука, 1981.

Красовский Н.Н. Управление динамической системой. Задача о минимуме гарантированного результата. М.: Наука, 1985.

Красовский Н.Н., Субботин А.И. Позиционные дифференциальные игры. М.: Наука, 1974.

Субботин А.И., Ченцов А.Г. Оптимизация гарантий в задачах управления. М.: Наука, 1981. 288 с.

Клейменов А.Ф. Неантагонистические позиционные дифференциальные игры. Екатеринбург: Наука, 1993.

Петросян Л.А. О новых сильно динамически устойчивых принципах оптимальности в кооперативных дифференциальных играх // Труды математического института им. Стеклова "Оптимальное управление и дифференциальные уравнения". Москва. 1995. Т. 211. С. 370–376.

Петросян Л.А. Характеристические функции в кооперативных дифференциальных играх // Вестник СПбГУ, сер. 1: Математика, механика, астрономия, 1995. № 1. С. 48–52.
Петросян Л.А., Данилов Н.Н. Устойчивость решений неантагонистических дифференциальных игр с трансферабельными выигрышами // Вестник ЛГУ. Серия 1, 1979. № 1. С. 52–59.
Yeung D.W.K., Petrosjan L.A. Cooperative Stochastic Differential Games. New-York, Heidelberg, London: Springer, 2006. 242 P.